ASSESSMENT OF THE PROCESS AND IMPACT **OF OPERATIONS RESEARCH** IN GUATEMALA: 1988 – 2000

Jane T. Bertrand
M. Celeste Marin

FRONTIERS/Tulane University

August 20, 2001

This review was funded by the U.S. Agency for International Development (USAID) under the terms of Cooperative Agreement Number HRN-A-00-98-00012-00. The opinions expressed herein are those of the authors and do not necessarily reflect the view of USAID.

Acknowledgements

We wish to thank those who facilitated the data collection and report writing for this assessment. Dr. Marcelo Castrillo and Dr. Carlos Brambila of the Population Council/Guatemala were very helpful in providing an overview of the operations research program in Guatemala, and Dr. Berta Taracena accompanied the evaluation team to various site visits. Jennifer Wheeler proved to be a very able assistant on logistics and support for this assignment.

We also thank Kris Lantis and Sarah Douglass of the Population Council/Washington for their editorial assistance and Kim Longfield, FRONTIERS/Tulane, for her assistance in compiling information for this report. Finally, we are appreciative to all who participated in the interview process. A complete list of persons interviewed is included in Appendix B.

This work was conducted under Cooperative Agreement No. HRN-A-00-98-00012-00. The opinions expressed are those of the authors and do not necessarily reflect the views of USAID.

Executive Summary

Rationale and Objectives

Population Council/Guatemala and the FRONTIERS Regional Associate Director for Latin America proposed this evaluation to document the utilization of the results from the portfolio of OR projects conducted over the past decade in Guatemala.

The objectives of the evaluation were:

- To determine the impact of the OR activities conducted on reproductive health services during this 12 year period, given the substantial level of project activity in Guatemala;

- To provide documentation of accomplishments and shortcomings of this OR work as the current cooperative agreement comes to a close;

- To identify factors that have facilitated and hindered either conduct or utilization of OR findings; and

- To test recent modifications to the FRONTIERS evaluation methodology.

Methodology

The evaluation team consisted of two FRONTIERS/Tulane staff: Jane Bertrand, Professor in the Department of International Health and Development, Tulane University, and Celeste Marin, Evaluation Specialist based in Washington, D.C. Both have worked closely with the evaluation methodology developed under the FRONTIERS Program. Data collection took place in April and May 2001.

The evaluation team relied on three primary sources of data: key informant interviews, document review, and several site visits to health centers and NGOs that have implemented OR interventions. Based on this information, they scored each of the 22 OR projects on 25 process and impact indicators, on a scale of one to three (3 being the highest score). They also collected data on contextual factors, but did not score them.

Results and Discussion

Projects reviewed were largely intervention studies, but included diagnostic and demonstration/evaluative studies as well. Table 1 gives an overview of study types and topics for OR projects during this time period, including those not part of this review. Individual scores for all indicators are presented in Table 2. In addition, the results section includes findings on each indicator, specifically the number of projects achieving the highest score, as well as a discussion of both high-and low-scoring studies on that item. The portfolio of studies performed best on the following items (with the proportion of projects receiving the top score given in parentheses):

- Research was relevant to local program managers (1.00).
- Results were disseminated to key audiences (.91).
- TA was provided in a sound and collegial manner (.89).
- Implementing organization participated actively in the design of the study (.86).
- Implementing organization participated actively in the conduct of the study (.86).
- Study was completed without delays that would compromise validity (.86).
- Results were judged to be credible and valid (.82).
- (If the intervention was effective and continued after the study) the activities tested under the intervention were still observable 36 months post-implementation (.82).
- Continuity in key personnel was maintained (.77).

Table 3 outlines the key findings from the 22 OR studies and indicates the actions taken by the implementing organization based on these results. In 13 of 20 studies (excluding the two that were diagnostic), the intervention proved effective. Four of the studies yielded mixed results, and three interventions were judged not to be effective. In 14 of the 22 studies, the implementing agency acted on the results. The successful intervention was scaled up within the organization in 9 of 17 projects, and adopted by another in-country organization in 5 of 17 projects.[1] Indeed, two USAID-funded projects in Guatemala – Calidad en Salud (URC) and NGO Strengthening (Population Council) – were designed with the expectation that they would take advantage of the lessons learned from this series of OR projects. Research findings did not generate substantial new funding, and project activities were seldom replicated in other countries (presumably because this was not a priority activity, nor were funds dedicated to doing so). However, a number of studies did lead to policy changes, primarily at the program level.

The assessment revealed several areas for improvement. First, several of the studies had too many objectives (making it difficult to fully achieve them) or inappropriate objectives (that described activities to be carried out, not results to be achieved). Second, most organizations did not build sufficient technical capacity to enable them to conduct subsequent OR projects. This finding underscores a dilemma for the Population Council: on one hand, they want to foster maximum skills-building and ownership of results in the implementing agencies; on the other, they are responsible to the donor agency for ensuring quality control at each phase of the research process. As such, they often "step in" to ensure a quality product, but in doing so they may defeat their own efforts at capacity building.

Some important factors in successful OR activities were identified as:

- Charismatic leadership, either from the implementing organization or the research team;
- Close monitoring and supervision of the intervention;

[1] The denominators in this section vary, depending on the number of studies that were applicable to the question. For example, some questions on impact were valid for all 22 studies, whereas others were relevant only for the 17 studies in which the intervention was effective and continued after the study.

- Simple, easy to use materials; and
- A feasible design.

Factors considered to increase utilization were:

- A good match between the intervention and the implementing organization;
- Immediate, observable improvements as a result of the intervention;
- Provider motivation;
- Continuing TA; and
- Fortuitous timing.

Diagnostic studies have been the center of debate on two points: (1) what is their value to an OR program, and (2) should they be assessed using the current methodology? We reviewed two diagnostic studies; two others that were labeled "diagnostic" on the original list were actually demonstration/evaluative studies. Although not all the indicators are applicable to diagnostic studies, this methodology can be used to track utilization of results from diagnostic studies.

Appendices
Appendix A contains summaries of the 22 OR projects reviewed. Appendix B lists the key informants interviewed for the evaluation and Appendix C contains the interview guide/data reporting form. A discussion of three proposed dissemination indicators is presented in Appendix D.

I. RATIONALE FOR THE EVALUATION OF OPERATIONS RESEARCH IN GUATEMALA

Two events prompted this assessment of operations research (OR) projects in Guatemala from 1988 to 2000. First, the Population Council/Guatemala is scheduled to complete a seven year cooperative agreement with USAID/Guatemala in December 2001, designed to investigate more effective ways of reaching the Guatemala population with reproductive heath services, particularly in Mayan areas. Second, the FRONTIERS Program has developed a methodology over the past two years to evaluate process and to document utilization of OR results. This methodology has been tested in six countries and is now available for use in relation to other portfolios of OR projects.[2]

The Regional Associate Director for FRONTIERS, based in Mexico City, and PC/Guatemala staff were interested in documenting the utilization of the results from the portfolio of OR projects conducted over the past decade in Guatemala, and they proposed this assessment to PC/Washington. (The types of studies and the substantive focuss the projects addressed are presented in Table 1.) The decision was then made for FRONTIERS/Tulane staff to carry out this assignment, given their familiarity with the evaluation methodology and with reproductive health services in Guatemala.

The Population Council's OR activity in Guatemala began in 1988, under the INOPAL project. At this time, INOPAL's regional office was located in Mexico City, and individual staff traveled to Central America to provide technical assistance on operations research to different service delivery organizations. Under INOPAL I, II and III a total of approximately 15 OR projects were conducted in Guatemala between 1988 and 1998.

In 1994, USAID/Guatemala awarded a cooperative agreement to the Population Council to establish an office in Guatemala City for the purposes of further developing a series of operations research projects with the Ministry of Health and local non-governmental organizations (NGOs) under Cooperative Agreement number 520-0357-A-00-4169-00. The objectives of this agreement were:

1. To develop new research-based strategies for the delivery of cost-effective reproductive health services to rural, Mayan and poor segments of the Guatemalan population;
2. To disseminate and promote the institutionalization of research findings to improve the quality of reproductive health services and to make these services more acceptable and accessible to the Mayan population;
3. To foster inter-institutional collaboration and cooperation;
4. To train Mayan professionals and support the integration of Mayan personnel into the field of reproductive health;
5. To improve the quality and coverage of maternal and child health (MCH) services provided by NGOs in rural areas, especially the Mayan highlands; and

[2] Bertrand, Jane T. and M. Celeste Marin, 2001. "Operations Research: Measuring Its Impact on Service Delivery and Policy," The FRONTIERS Program, Washington, DC.

6. To improve the management and sustainability of MCH programs carried out by NGOs.

Under this cooperative agreement the Population Council, in collaboration with local implementing organizations, conducted a total of 22 OR projects (defined as having a distinct project number) between 1994 and 2001. This review includes those projects completed by 2000 to allow sufficient time for results to be utilized. (See Appendix A for project summaries.) The Population Council is also involved in five other OR projects in Guatemala under FRONTIERS as well as eight under the cooperative agreement, but these have yet to be completed and are not part of this review.

Guatemala constitutes an excellent site for operations research since the delivery of family planning and reproductive health (FP/RH) services in this country still trails that of neighboring countries. Guatemala has the second lowest contraceptive prevalence rate (38 percent) of any Latin American country; the only lower is Haiti. Whereas 50% of Ladino women married or in union report using contraception, this is true of only 13% of Mayans. Maternal mortality is also high by regional standards: 190 deaths per 100,000 live births,[3] with marked disparity by ethnic group. Various factors explain this large gap between Mayans and Ladinos (which exists on almost every possible social indicator). Three-quarter of Mayans live in poverty. The highlands of Guatemala (home to many Mayans) were ravaged by the civil war that reached its peak in the 1980s, crippling many of the social programs in that area. And the Catholic Church has exerted tremendous pressure on the government to curtail family planning services over the various administrations of the past 30 years. Despite this sobering backdrop, progress has been made in improving service delivery and in reaching out to the Mayans with culturally appropriate strategies.

The primary focus on this evaluation concerns the utilization of results to improve service delivery and influence policy (referred to herein as "impact").[4] Although we were interested in the quality of research (and measured some aspects of it using the process indicators), our main task was to assess the extent to which these OR studies resulted in observable change in the service delivery environment for family planning and reproductive health.

This evaluation represents a welcome opportunity to apply the methodology previously developed under FRONTIERS in a country with substantial OR activity. This report contains the major findings from the evaluation, as well as observations on the methodology itself.

[3] World Bank, *World Development Indicators 2000*.

[4] We use the term "impact" with some reservation. To meet the test of measuring impact in the rigorous sense of the word, the evaluator needs to be able to demonstrate cause and effect. He/she should be able to show what would have happened in the absence of the program or project. Although we do not meet this standard with the current methodology, we have retained the word impact because it captures the sense that "something changed" in the aftermath of the OR study.

II. METHODOLOGY

A. Objectives of the evaluation

The objectives of this evaluation were:

- To determine the impact of the OR activities conducted on reproductive health services during this 12 year period, given the substantial level of project activity in Guatemala;
- To provide documentation of accomplishments and shortcomings of this OR work as the current cooperative agreement comes to a close;
- To identify factors that have facilitated and hindered either conduct or utilization of OR findings; and
- To test recent modifications to the FRONTIERS evaluation methodology.

B. Composition of evaluation team

Jane Bertrand and Celeste Marin visited Guatemala from April 2-6, 2001, and Marin returned from May 21-26, 2001 to collect data for this evaluation. Both members of the evaluation team are FRONTIERS/Tulane University evaluation staff. Jane Bertrand is Professor in the Department of International Health and Development, School of Public Health and Tropical Medicine. Celeste Marin is Evaluation Specialist under FRONTIERS, based in Washington, D.C. Both have worked closely with the evaluation methodology developed under the FRONTIERS Program, but had no previous involvement in any of the projects in this review.

C. Sources of data

The evaluation team relied on three primary sources of data: key informant interviews, document review, and several site visits to health centers and NGOs that have implemented OR interventions.

In April 2001 the team conducted key informant interviews in Guatemala City with Population Council and collaborating researchers, administrators and managers from implementing institutions who were involved in the study, representatives from USAID/Guatemala, and "potential users" — individuals in a position to utilize OR results. The team also interviewed providers in Sololá. In May 2001 Marin conducted additional interviews on site in Quetzaltenango, San Marcos, and Totonicapán with area administrators, service providers, and NGO staff, including rural health promoters. A list of all persons interviewed is presented in Appendix B.

The evaluation team used the Assessment Form (see Appendix C) as an interview guide. Although most key informants were able to answer only some of the questions from the assessment form, the team interviewed several individuals per project, which yielded sufficient information to complete all grids on almost all indicators.

In addition to interviews the team reviewed project reports and other documents on the design and implementation of the studies. Each study had a final report in English or Spanish, and most were summarized in publications such as "Findings and Lessons Learned in Delivery of Reproductive Health Care to the Rural Mayan Population of Guatemala from Operations Research and Diagnostic Studies, 1994-1997,"[5] *El Pregón*, a magazine aimed at NGOs, or brief handbooks for program staff or government officials describing the merits of a study or how to replicate the intervention. The final report in particular was used to evaluate the studies on the indicator P-9, "Was the study design methodologically sound?"

D. Limitations of the methodology

There are several limitations to this approach to assessing impact of OR projects. First is the issue of attribution. Other organizations in Guatemala have been conducting related activities, and events may have occurred concurrently that encouraged service delivery organizations to adopt changes, independent of the OR study. Thus, it is virtually impossible to demonstrate cause and effect in this type of assessment of impact of OR studies. Rather, we seek to demonstrate "plausible attribution," which requires that 1) those implementing the new procedure or approach know of the OR results, 2) the change in service delivery take place after the intervention, and 3) the change that occurs is consistent with the results and recommendations of the OR study.

Second, turnover of staff in government and NGOs and physical inaccessibility made it difficult to locate some individuals who participated in a study or were potential users of the findings. In some cases, those who had participated in the project could not be found, and their replacements had very little knowledge of project activities. We interviewed a minimum of two informants for each project, but in some studies we gathered information from more sources than usual to minimize the effect of not being able to contact a key person. This problem resulted primarily from the lapse in time between the completion of many of these studies and this review. It is not an inherent flaw in the methodology and should not be a problem when the methodology is applied prospectively to the portfolio of FRONTIERS projects.

Third, the assessment of process and effect is qualitative in nature. The evaluation team must gather information from various sources and make subjective judgments in rating a specific project on a specific indicator. However, the approach is systematic; a set of

[5] Population Council/Guatemala. 1998. *Findings and Lessons Learned in Delivery of Reproductive Health Care to the Rural Mayan Population of Guatemala from Operations Research and Diagnostic Studies, 1994-1997.* New York: Population Council.

indicators was applied to each subproject using a predetermined discussion guide. When the responses of various informants were combined they usually supported one another, and, in cases where they were contradictory, additional information was gathered to determine how to rate the project on that indicator.

III. RESULTS

A. Projects reviewed

A complete list of all projects included in this review appears in Tables 2 and 3. It proved important to differentiate between the different types of studies (see column 2 of Table 2). The types of studies, from least to most "rigorous," are as follows:

- Technical assistance (0) [6]
- Diagnostic (2)
- Demonstration/Evaluative (9)
- Intervention (11)

Although the lines blur between the different types of studies, we defined the study types as follows:

Technical assistance: attempted to strengthen the different functional areas (supervision, training, IEC/counseling, MIS) with the aim of improving service delivery in previously established programs.

Diagnostic: consisted of research for the purposes of better understanding the target population, their motivations and fears, their language and perceptions, and other factors relevant to their potential acceptance of a given reproductive health intervention.

Demonstration/Evaluative: involved some collection of data to measure trends in utilization, client satisfaction, or other behaviors, with the intention of documenting results or identifying areas in need of improvement. The term "demonstration" is often used when a new service is offered for the first time.

Intervention: involved some type of experimental or quasi-experimental design to evaluate the effectiveness of the treatment against a second strategy or the status quo.

The methodology for evaluating OR projects is most appropriate for the last two types (demonstration/evaluative and intervention) because the impact questions on scaling up, replication and so forth are relevant. At the request of the Population Council, we also included diagnostic studies. Given that they can serve to "test the waters" for the potential acceptability of new reproductive health interventions, it is appropriate to ask, "What happened in terms of service delivery as a result of these diagnostic studies?"

[6] This category refers to projects that are exclusively technical assistance, without any research component. The vast majority of the 22 projects included some aspects of TA, but it was in addition to the research component. The PC representative for project #10 described this project as being "basically TA," but we chose to label it as demonstration/evaluative, based on the criteria listed in this section.

B. Format for presenting the results

The results of this assessment are presented in several ways. First, the overall result for each indicator (expressed as the number of projects of the total reviewed that receive the highest score [3]) is presented as a bullet, in bold. Although this review covers 22 studies, the denominator for "total number of studies" drops as low as 17 in those cases where the question is not applicable (e.g., the items on scaling up interventions were not applicable to diagnostic studies).

Each project is scored on each indicator using a scale of one to three, with the numbers corresponding to the following definitions:

1 — slightly or not at all (up to 1/3 of potential)

2 — somewhat (1/3 to 2/3 of potential)

3 — a great deal (2/3 to full potential)

This scale distinguishes between those studies that performed well (3), those that performed satisfactorily but with notable problems (2), and those that did not perform satisfactorily on the relevant indicator (1).

The indicators fall into two main categories: process and impact. Within these two main categories certain indicators cluster naturally, as indicated by the underlined subtitles in the section below. The numbers in parentheses (e.g., "P-4," "I-5") refer to the numbers of the process and impact indicators, respectively, as they appear on the data collection instrument, shown in Appendix C.

We also collected data on six contextual factors, but rather than presenting them in the same format as the process and impact indicators, we have woven them into the text where they are relevant. The contextual factors are not used to "rate" a project but rather to understand what factors beyond the control of the program managers and researchers affect the study and utilization of its results.

Second, after each bullet we have provided a more qualitative assessment of the point, bringing to light material that emerged during the interviews with key informants.

Third, Table 2 provides a tabulation of results for all 22 studies across the 14 process indicators and 11 impact indicators.[7] This allows one to assess the overall performance of each study (the "perfect score" being a column of 3s) and the performance of the

[7] Note: there is no P-13. Whereas one might assume this reflects the North American aversion to the number 13 as unlucky, in fact it is omitted for a different reason. In previous versions of the instrument, it referred to whether a study included a cost analysis. We subsequently realized that its inclusion in the grid implied that all studies SHOULD include a cost component, which is not the case. Thus, we moved the item related to cost to the section on "contextual and other factors," to avoid its counting on this score sheet that some will interpret as a report card. We have retained the original item numbers so the indicators would be consistent with the numbering systems in other OR evaluations.

portfolio of studies on specific items (the perfect score being a row of 3s for that indicator). The final columns on Table 2 show the number of studies scoring a three over the total number of applicable studies for that item; this is translated into a proportion of studies scoring 3 in the final column, with a possible range of 0.00 (none) to 1.00 (all).

The numbering of the projects from 1 to 22 was done (1) for the convenience of the evaluation team and readers of this report, and (2) to establish the sequence of projects conducted by a given organization (e.g., APROFAM: #1-6, AGES: #7-9). It does not correspond to any Population Council numbering system.

C. Results for process indicators

Collaboration between the Population Council and implementing agencies

The main implementing agencies were the Ministry of Public Health and Social Assistance (referred to hereafter as "MOH") (n=5 studies), the Guatemala Social Security Institute (IGSS) (n=2), Asociación Pro-Bienestar de Familia (APROFAM) (n=5), and Guatemalan Sex Education Association (AGES) (n=3). In addition, the following organizations carried out one OR study each with the Population Council: Asociación Pro Salud Preventiva para la Mujer Vivamos Mejor (APROVIME); AGROSALUD; Asociación Toto-Integrado (ATI); Promoción, Investigación y Educación en el Salud (PIES del Occidente)/Coopreración para el Desarrollo Rural de Occidente (CDRO); Project Concern International (PCI); Rxiin Tnamet and Universidad del Valle.

- **In 19 of the 22 studies, the implementing organizations actively participated in the design of the study (P-1).**

This group of studies got high marks for the active collaboration of implementing organizations in the design of the study. Many key informants spoke of the animated discussions that they had with Population Council staff in identifying priority topics and deciding which to select for an OR study.

However, on more than one occasion their participation was limited to agreeing to the design proposed by the Population Council. Whereas the implementing organization clearly recognized the relevance of the topic to a problem they faced, it was often the Population Council staff and not the organization per se that identified the problem. We revisit the issue of the appropriate role for the technical assistance organization in the discussion section below.

- **In 19 of the 22 studies, the implementing agencies actively participated in the implementation of the study (P-2).**

As a group, the implementing organizations also reported a high level of participation in the actual conduct of the study. This included one or more aspects of designing the questionnaire, recruiting the field personnel, pre-testing the instrument, collecting the data, cleaning and processing the data, analyzing the data, and writing the report.

However, the "high score" on this indicator masks the fact that many organizations participated in some but not all of these tasks. The Population Council conducted and managed the MOH and IGSS studies, as these organizations had little to no experience in operations research for reproductive health. APROFAM, an organization with OR experience and research personnel, often found it difficult to free up its regular employees to work on its OR projects and instead hired consultants to oversee the conduct of the research. This strategy was very effective in completing the projects but resulted in more limited "ownership" by APROFAM once the study was done.

Despite participation in specific parts of the project implementation, the implementing organizations did not necessarily increase their capacity to do operations research; field supervisors, coordinators and analysts were generally Population Council staff or consultants, and local level staff carrying out the intervention did so with training and assistance from the Population Council.

- **In 13 of the 21 studies, the implementing organization actively participated in developing programmatic recommendations (P-6).**

Recommendations were usually developed in the course of writing the final report. If Population Council staff alone wrote the report, they proposed recommendations and solicited feedback from their counterparts. In these situations the implementing organization had some input on the final conclusions and recommendations, albeit less than the Population Council. In other instances Population Council and the implementing organization jointly prepared the final report, leading to more balanced participation.

There were two important exceptions to the above patterns. In the first case, a representative from the National Nursing School approached the Population Council for assistance in developing a distance education program (#20). The Nursing School had a major role in designing and implementing the intervention and in developing the recommendations. In a second case, an anthropologist from a local university received terms of reference (and little more) from the Population Council for the study of cognition and speech patterns concerning reproductive health among Mam speakers in Quetzaltenango (#3). He and his staff designed and conducted the study and prepared the final report, complete with conclusions and recommendations, with only logistical assistance from the Population Council.

- **In 17 of the 19 studies, the counterparts in the implementing agencies judged that the technical assistance was useful, methodologically sound, and provided in a collegial manner (P-8).**

The key informants were uniformly positive, even glowing, in their praise of the technical assistance that the Population Council provided on the OR studies. They felt it was both methodologically sound and delivered in a very collegial manner. Many spoke with great affection of the different Population Council staff members that had worked with them on the different studies. Negative comments were rare. In one case, previously cordial relations seemed to sour as the project ran into insurmountable problems and was not able to achieve its objectives (#9). In a second case, different key informants remembered the attitude of the consultant in slightly different terms. Note: three studies were excluded on this indicator: one because it did not receive technical assistance from the Population Council (#13), and two because we were unable to obtain the information from the appropriate individuals (#12 and #14).

Quality of study design and implementation

- **In 12 of the 21 studies, the study design was judged to be methodologically sound (free of flaws that could have affected the final results) (P-9).**

On the whole, the studies evaluated were methodologically sound. Indeed, no study got a score as low as "1." Rather, a number of the studies only scored a "2" (indicating some problems on this indicator), due to data quality, loss to follow-up and designs that were not sufficiently scientifically rigorous for this type of research. For example, a diagnostic study of TBAs' knowledge (#16) was conducted prior to developing a training curriculum, but acquired knowledge was measured using a post-test only design with a separate sample. Another threat to some designs was that service statistics and other existing data to be used were not always available or reliable, so either alternate strategies had to be developed (with mixed success) or certain objectives could not be achieved or measured. Such was the case in the TBA study (#16) and the tripartite strategy in San Marcos (#21).

- **In 13 of the 20 studies, the intervention was implemented as planned (or with some modifications) (P-4).**

Most interventions were implemented as planned, but there were a few notable exceptions. The AGES project on "Access to Reproductive Health Services and Education in Indigenous Communities" (#9) ran into serious problems in introducing a service delivery component. Similarly, the APROFAM project to introduce reproductive health education to men (#3) proved particularly challenging, and the original strategy involving formal health education talks was abandoned in favor of more recreational and athletic activities. The ATI-APROFAM study (#14) did not even make it beyond the diagnostic stage; the two NGOs spent a great deal of time clarifying their relationship and

had major funding problems, so the study ended early and the intervention was not implemented at all.

- **In 19 of the 22 cases, the study was completed without delays that would compromise the validity of the research design (P-5).**

Many studies had delays but these tended to occur in the start up. In most cases the end date was simply pushed back to accommodate the intervention and the delay had no impact on the design or results. In others, such adjustments were not possible. The proposal for the distance education course (#20) had to be rewritten to gain approval and, as a result, the final module of the course could not be completed before the study ended. The reproductive health educational strategies for men (#3) experienced such difficulty identifying activities that would attract men that they had insufficient time to conduct and evaluate these activities. Not surprisingly, the evaluation showed little effect of this strategy on client behavior.

- **In 17 of the 22 studies, there was continuity in key personnel over the life of the project (P-7).**

The large majority of studies enjoyed continuity of key personnel over the life of the project. However, there were some exceptions. The APROFAM project on re-engineering the CBD program (#5) experienced significant difficulties with turnover not only of the principal investigator, but also the executive director of the organization and other key administrative personnel.

In other cases, key personnel were maintained throughout the studies but not necessarily beyond completion, limiting utilization. For example, in the TA activity "Strengthening Knowledge and Management Skills for MOH Family Planning National Supervisors,"[8] all trained supervisors were transferred, fired or resigned immediately following the training, leaving absolutely no impact.

Although this indicator is worded to imply that continuity is positive, in a few cases a change in key personnel proved to be an advantage. For example, the current MOH administration is much more supportive of family planning and reproductive health than the previous one, and more committed to improving quality and access to services.

- **In 13 of the 22 cases, the study accomplished the research objectives (P-3).**

A common problem in a number of the OR studies was an overly ambitious research agenda. Some of the research projects had up to ten objectives, making it difficult to realize them in full (see discussion section below for additional observations on the

[8] This project was riginally included in this evaluation but later removed because it was a TA activity and not an operations research study.

achievement of objectives for these OR projects). Only two of the 22 studies (#9, the AGES study on Access to Reproductive Health Services and Education in Indigenous Communities, and #14, Collaboration between Two NGOs, ATI and APROFAM in the Delivery of Family Planning Services) fell entirely short of the mark, and low scores on this indicator were associated with low scores on P-4, "The intervention was implemented as planned." Had the objectives been less ambitious, it is likely that a larger percentage of the studies would have received a top score on this indicator.

Appropriateness to local context

- **In 15 of the 22 studies, the research design was feasible in the local context (P-10).**

This seemingly straightforward indicator proved multifaceted in its application. "Feasible" can refer to either the study design or to the actual implementation. If it refers to the design, it relates to the previous indicator about achieving the research objectives. For example, an early APROFAM study (#1) on testing different training and supervision strategies among Mayan communities proved to be more challenging and placed a greater research burden on the service delivery organization APROFAM than they had anticipated. It proved difficult to avoid contamination between the experimental and control groups, given program dynamics.

Examples related to implementation include the following. (1) Many of the IGSS-affiliated TBAs (#16) were illiterate and could not keep service statistics. An alternate strategy to track referrals by issuing clients cards did not work, so no conclusions could be reached on referral rates. (2) The IMCI study (#22) struggled with an overly complex analysis plan. (3) The ATI-APROFAM (#14) collaboration suffered from internal problems (interpersonal conflicts and miscommunication between the two organizations) as well as external difficulties, including insufficient funding for both NGOs. While neither internal problem was caused by the study design, they might have been avoided by allowing more time at the beginning of the study to develop the relationship between the two NGOs.

In short, factors that caused the research design not to be feasible included: (1) the need for rigorous controls in studies with a quasi-experimental design, (2) data collection by illiterate health workers, and (3) data analysis that exceeded the capacity of those expected to do it.

- **In 22 of the 22 studies, the research was judged relevant for local program managers (P-12).**

All studies were relevant to the priority populations for this Cooperative Agreement: indigenous groups in the western highlands or (in the case of IGSS) workers employed in the formal sector. The studies tested proposed solutions to important problems or

facilitated the later development of appropriate solutions through formative research. Existing reproductive health services were modified to provide a broader spectrum or better quality of care and, in situations where reproductive health services were not yet offered (such as on agricultural farms or in postpartum/postabortion wards), they were incorporated into existing services, taking advantage of the infrastructure available.

- **In 18 of the 22 studies, the results were judged to be credible/valid in the local context (P-11).**

Virtually all of the interviewed informants — including researchers, administrators and providers — believed that the results and recommendations that they knew about were credible and valid. Although many knew of results only from a few specific studies, they regarded all Population Council activities to be of high quality and judged Population Council presentations and publications to be reliable sources of information. (In a previous report, we referred to this as a "halo effect").

Several of the informants did say that "some providers" were not so convinced by the study results and did not want to follow the recommendations. In particular, these other providers did not use the MOH algorithm because it took up too much time and they were unable to see any benefit in offering women services they were not seeking. However, the "key audiences" (policymakers, program managers and other decisionmakers) were unanimous in their high opinion of the credibility and validity of OR results from the Population Council.

In two of the studies that received a "2," key informants questioned whether the results had even been diffused, making it difficult to give this project a top score. In another case (#13), a senior Population Council staff member expressed strong doubts about the recruitment of the study population and therefore the validity of the results. The PI also presented the results recently at an international conference on men in reproductive health in Buenos Aires to a mixed response, which he attributed to the incompatibility of his findings (that Mam men associate sex with sadness and guilt, among other things) with the more widely held views on "machismo."

Dissemination

- **In 20 of the 22 studies, results were disseminated to key audiences, including policymakers, program managers, and service providers (P-14).**

The studies had no uniform dissemination strategy, but in nearly all cases the Population Council used more than one approach to reach key audiences. The minimum requirement of all studies was a final report, but for most studies the Population Council also produced a summary of results or a booklet containing guidelines for replicating the intervention. In this way different audiences received results in different formats according to their needs. In 2000 PC/Guatemala began publishing *El Pregón*, a magazine for the NGO community. This magazine contains articles about reproductive health research and activities, as well as news on related topics such as adolescents, community

development and violence. It is distributed to a number of NGOs directly and to others through international organizations. Project Concern International (PCI), for example, distributes copies to the NGOs with whom it collaborates and praised *El Pregón*'s content and simplicity. Population Council staff and their research partners also disseminate research findings through presentations. Results are usually presented first to local staff that participated in the intervention and then at the central level, often at the collaborating institution as well as the Population Council office. The Population Council also convenes an interagency Reproductive Health Working Group that meets regularly, providing a forum for such information exchange among USAID, CAs and NGOs working in the reproductive health field.

Different research projects target different "key audiences" for utilization. In the MOH and IGSS, for example, only one or two key individuals need to be convinced of the merit of the research, since they are the main reproductive health decisionmakers. Because these key people are also involved to some degree from the beginning of the study, dissemination takes place throughout the entire process, not only at the end. In other studies, for example the diagnostic study of the Mam, much more dissemination is needed as those who are in a position to utilize the findings provide services in other organizations. (In this case the final report, several summaries, and presentations in three locations proved inadequate.)

Informants made several comments about shortcomings in dissemination. One IGSS representative said that they never saw a final report or summary of the TBA study, which was completed four years ago. The only report prepared was in English, which was of limited use to them. It was not clear whether they had ever addressed this with the Population Council, but the informant did ask the evaluation team for a Spanish summary of the results because people continue to ask him about the study. Another informant complained that in general, research is not presented at an appropriate level for users. He said the Population Council was no exception and should invest more effort in simplifying dissemination materials and reaching wider audiences. Although he does not currently collaborate with the Population Council, he is a key figure in the reproductive health community.

While the above media have proven effective for promoting utilization, the Population Council has the benefit of two other USAID-supported mechanisms: *Calidad en Salud* and the NGO strengthening component of the cooperative agreement.

Calidad is a consortium of reproductive health organizations that provides technical assistance to the Ministry of Health. Two Population Council staff seconded to Calidad played important roles in OR at the Population Council, and the organization relies on many Population Council materials and lessons learned in its work. According to one, "We are now reaping the benefits of our years of hard work and research." The Calidad-Population Council link has accelerated the pace of utilization of OR findings and OR-generated materials such as job aids, training manuals and IEC materials. While Calidad is not adopting all Population Council interventions as is (some aspects must be modified

to apply to the whole country, not just the western highlands), OR products and results do form the basis of much of the TA provided.

The NGO strengthening component functions in a similar but more direct manner. The Population Council itself provides the TA, and NGOs — being smaller and more independent than a large institution like the MOH — can implement changes more easily. Some NGOs (CDRO, B'elejeb B'atz, PIES del Occidente) have participated in an OR study. Others have not conducted OR but have used OR results to improve their own services after seeing the effects on other organizations. The continuing interaction inherent to the Population Council-NGO relationship promotes not only adoption but also institutionalization of changes, while maintaining enough flexibility to ensure that the intervention remains a good fit for the organization. The NGO strengthening component allows the Population Council to pair operations research and technical assistance to maximize the impact of both.

- **In 15 of the 22 studies, results were readily available in written form to interested audiences (P-15).**

The scoring on this item was somewhat arbitrary but used the following rule: if the results were included in a publication that received wide circulation, the study got 3 points (the highest score possible); if the results existed only in the form of a final report, it received 2; and if there was no report on file it received 1 (though no project scored a "1" on this indicator).

D. Results for impact indicators

Intervention effective, acted on

- **In 13 of the 20 studies, the results indicated that the intervention was effective (i.e., that it improved service delivery in areas identified by the OR study) (I-1).**

The indicator I-1, "the results indicated that the intervention was effective," was primarily intended for intervention studies (of which there were 12). However, we have broadened it to include demonstration/evaluative studies, which allowed us to review all but two (the "purely diagnostic" studies).

Table 3 summarizes the key findings from all 22 studies reviewed. In the 20 studies involving an intervention, 13 showed the intervention to be effective (i.e., as a result of the intervention the project achieved its stated objectives). Based on a purely qualitative assessment, the following six interventions that stood out as "most effective" (listed here by number, with a one-line summary of the main finding):

#4 Injectables: Depo-Provera can be distributed through CBD promoters at the community level.

#6 Referrals to APROFAM clinics: **The strategy of bringing promoters to the clinic and training them onsite is more effective than the status quo training system in**

increasing the number of referrals they make to APROFAM clinics, but it also is more expensive than the status quo.

#11 Necklace: The necklace method for facilitating the use of rhythm was both effective in preventing pregnancy and acceptable to Mayan couples who elected it.

#15 Postabortion care: Many women not using family planning will adopt a method if services are more convenient (in particular if they can get a method right away).

#19 Algorithm: Offering integrated reproductive health and infant services systematically through a standard set of questions asked of all women regardless of the reason for visit improves quality of care and client satisfaction.

#21 Re-engineering of MOH clinics: When providers feel they have the ability to make changes and can see the benefits, they are willing to work harder to provide quality care; in turn, satisfaction of both clients and providers increases.

Four of the 20 studies yielded mixed results — the intervention was not effective as such, but some part of the study was adopted and used to improve service delivery. For example, the first algorithm study (#17) showed that in several cases the control group outperformed the experimental group, but those doctors who used the algorithm regularly in their consultations (some of the experimental group doctors did not) consistently outperformed the others in both groups. As a result, the algorithm was seen to be a potentially effective tool and the Population Council and MOH collaborated to modify and test it in subsequent studies. In a second case, training TBAs (#16) did not prove to be a cost-effective strategy for improving the quality of care for rural pregnant women in the IGSS system, but the researchers found the supervisory system introduced under the project reinforced the training and effectively maintained and improved quality. The study also demonstrated that including self-esteem in training for community workers could increase the learning and retention of knowledge and skills.

Of note, three of the 20 interventions were judged **not to be effective**. This underscores a very important point: not all interventions or strategies in an OR portfolio will "work." If we could assume a near 100% success rate, we would not need to conduct the studies themselves but rather we could direct all funds into implementing the interventions. However, it is exactly because "some will work, and some don't" that OR serves such an important role in providing evidence-based guidance to program managers who must select between options.

- **In 14 of the 22 studies, the implementing agencies acted on the results (I-2).**

This indicator applied to all 22 studies and measured whether the organizations under review took specific measures – consistent with the results of the study – to improve service delivery. We defined "acted on" as continuing to implement the activities tested in the OR study after its completion if effective, or **not** implementing or discontinuing the activity if ineffective.

The scores for I-2 (whether the implementing organization "acted on" the results) tended to correspond to the score for I-1 (whether the results indicated that the intervention was effective). The organization acted on the results in every case where the study received a top score on the effectiveness of the intervention, with one exception (improving client referrals to APROFAM clinics, #6). In that study, the lack of utilization related both to turnover in top-level personnel and the cost associated with the more effective strategy.

In the evaluative study on re-engineering the CBD program (#5), APROFAM took very concrete actions: reducing the number of distributors by eliminating some of the less productive ones and adopting a profile of the ideal CBD workers in Mayan areas (for future recruitment).

In the case of demonstration/evaluative projects, services introduced were maintained and in some cases scaled up (re-engineering of MOH health center services [#21] and FP counseling and methods for postpartum/postabortion women [#15]). The results of interventions that were only partially effective were acted on less completely. For example, the results of the TBA study showed that training could improve the knowledge and skills of TBAs, but IGSS chose not to act on this because even a large investment in training would have little impact at the population level; most TBAs attended about three IGSS-affiliated births per year, while the maternal hospital attended approximately 15,000. A small number of TBAs attended many births, but IGSS's equity policy did not allow for selecting only certain individuals for training.

- **In 14 of the 17 studies, (if the intervention was effective and continued after the study) the activities tested under the intervention were still observable 36 months post-implementation (I-3).**

Overall, those organizations that had acted on the results of an OR study tended to stick with the new strategy or intervention for a substantial period thereafter. (Note: although the item stated "36 months," we scored five interventions on this item that had been implemented more recently than 36 months because all evidence suggested the intervention was "there to stay"). The proportion of interventions that "stuck" is relatively high: 0.82.

Evidence of capacity building on OR

- **In 16 of the 22 studies, the implementing organizations conducted subsequent OR studies (I-4).**

Ideally, the experience of conducting an OR study should serve to increase the technical capacity of implementing agencies to subsequently use this technique in relation to other problems. The evaluation addressed this question in two forms: "Did the implementing agency conduct subsequent operations research?" and "Did the implementing agency conduct subsequent operations research without the assistant of the Population Council?" (the latter presumably being more difficult).

In 16 of 22 cases the implementing organizations conducted subsequent OR. This result masks the reality, however, that three organizations – the MOH, APROFAM, and AGES – each conducted at least two OR studies in this portfolio and thus are credited with "subsequent OR" when in fact they were simply conducting more studies under the same funding mechanism.

The response varied greatly by organization. For example, given that APROFAM has a long experience with applied research, it is not surprising that they received the top score across the board on this indicator. The MOH received the top score, in large part because of the subsequent rounds of testing for the algorithm with the Population Council. In contrast, the studies conducted with smaller NGOs or the local university researcher were a "one shot deal," in which the researcher or organization did not conduct subsequent operations research.

Both the MOH and IGSS had numerous problems in providing services and realized that working with the Population Council could help them solve these problems. As a result they have been willing to collaborate with the Population Council on numerous studies, continuing to present day, although neither institution seems to be interested in conducting OR without Population Council assistance. Both MOH and IGSS conduct some research, but it consists primarily of epidemiological research or analysis of service statistics.

- **In 8 of the 22 studies, the implementing organizations conducted subsequent OR studies without the Population Council (I-5).**

The pattern of subsequent operations research **without the assistance of the Population Council** was even more dramatic. The only organizations that answered affirmatively to this question were APROFAM, AGES, and APROVIME (staffed by previous APROFAM employees).

When asked whether IGSS would continue to do operations research without technical assistance — or even desire to — an informant from IGSS stated, "IGSS does not have a culture of research. Our main interest has always been in providing services." Similarly,

the Ministry of Public Health does not have or anticipate having a research unit or staff with the capacity to conduct OR but, unlike IGSS, it does work closely with a research NGO, the Center for Epidemiological Research in Reproductive (CIESAR). Recently the MOH and CIESAR have jointly collaborated with WHO and PAHO, and organized a regional PAC conference with the Population Council. It is possible that this group will continue to carry out research activities with the MOH as a parallel institution for the long-term, much as the Nutrition Institute for Central America and Panama (INCAP) conducts nutrition research with the MOH.

Scaling up and replication

- **In 10 of the 17 studies, (if the intervention was effective and continued after the study) the intervention was scaled-up by the original implementing organization in the same country (I-6).**

This item and the next two (I-7 and I-8) were asked of only 17 studies; the two diagnostic studies and the three interventions that were not shown to be effective were excluded.

Sixteen of the 17 projects were scaled up to some degree, but we assigned six of the projects a "2" instead of a "3" because we considered the scaling up to be limited. Of the ten projects more fully scaled up, examples that stand out include the distribution of injectables by CBD workers in APROFAM and the promotion of the necklace method by indigenous NGOs. Systematic offering of integrated services through use of the algorithm (#17-19) was scaled up gradually through the three studies in this evaluation, in the areas of Chimaltenango, Totonicapán, San Marcos, Quetzaltenango and Sololá. The Ministry of Health is now preparing for a complete, nationwide scale up. Scaling up of the Community IMCI intervention (#22) has been quicker. The Community IMCI is a modified version of MINE (Manejo Integrado del Niño Enfermo), which had been in use for several years prior. In addition, IMCI is a tool recommended by WHO and PAHO, giving it greater credibility than one developed by the Population Council alone. But perhaps the most important reason for the widespread use of IMCI is that it responds to the more urgent health needs of Mayan communities: acute childhood illnesses which service providers and community promoters must deal with regularly. In this context it has been important to improve not only reproductive health services but also primary care for children. As a result, community health promoters can respond to more needs of the populations they serve, increasing respect for their work. The Population Council has plans to combine the two algorithms into one MCH instrument in the future.

Technical assistance for re-engineering, the primary component of the tripartite strategy tested in San Marcos (#21), was available for those health centers choosing to participate immediately following the study. The intervention was also adopted by two other Health Areas, Sololá and Quetzaltenango, but only carried out in Sololá because Quetzaltenango's director was replaced, and the new director was not interested. Those centers in San Marcos that did not choose to participate immediately after the study but have since changed their minds have not been considered by the Health Area for scaling up, and continue to operate as they had in the past.

As a smaller institution with few hospitals, IGSS was able to scale up postpartum/ postabortion family planning services fairly easily (#15), and implement the new supervisory strategy system-wide (#16). While not yet implemented beyond the original area, the Nursing School's distance education project (#20) is available to new nurses entering the system, thus increasing the number participating in the course.

- **In 5 of the 17 studies, (if the intervention was effective and continued after the study) the intervention was adopted by another organization within the same country (I-7).**

Replication of OR interventions by other organizations has occurred almost exclusively through the NGO strengthening component of the cooperative agreement. The two algorithms, for FP/RH services and childhood illness, have been most applicable to the needs of these NGOs. They are used with differing levels of adherence — one NGO now requires prior experience using the reproductive health algorithm of all new doctors, while the other NGOs have a policy mandating its use but acknowledge that a number of their doctors don't use it at all, or do so irregularly. The NGOs that don't require doctors to be previously trained may have difficulty maintaining trained staff without financial assistance from the Population Council, as their own funds are very limited, and doctors leave frequently for more prestigious, urban jobs.

Calidad en Salud provides technical assistance to the Ministry of Health using materials and lessons learned from the Population Council OR (as well as from other organizations) but this cannot be considered as replication by another organization, as the Population Council is a member of the consortium.

- **In 0 of the 17 studies, (if the intervention was effective and continued after the study) the intervention was replicated in another country (I-8).**

No one interviewed in Guatemala was aware of any intervention being replicated outside of the country, but many acknowledged that even if an intervention had been replicated, they would not necessarily know about it. Those outside of Guatemala who had a broader view of reproductive health programs in the region suggested that two interventions in particular had been replicated elsewhere: use of the "necklace method" for fertility awareness and natural family planning (#11 and #12), and the algorithm for providing systematic integrated health care services for women and infants (#17 – 19).

The necklace is indeed being used or tested worldwide by the Population Council, local NGOs and Georgetown University's Institute for Reproductive Health (IRH). Because the Population Council first decided to test the effectiveness and acceptability of using a necklace to teach fertility awareness after reading about a necklace being used in Cote D'Ivoire, they cannot claim to be the originators of the tool. Nevertheless, the necklace component of the 1992 AGROSALUD study (#12) does predate any other research or use of the necklace in Guatemala by several years. Following the AGROSALUD study, Population Council/Brazil asked for the final report and went on to conduct an OR study on the necklace as a family planning tool. IRH reported basing their studies on a necklace

used in Brazil. We were not able to determine to what extent the Guatemala study influenced the Brazil study, nor whether the woman who told IRH of her experiences in Brazil had any connection with the Population Council study, but we consider it plausible that the AGROSALUD study did impact later work within and outside of Guatemala.

PC/Guatemala and PC/Mexico (who had coordinated the AGROSALUD study before the Guatemala office opened) collaborated in designing the 1997 APROVIME study. The Deputy Director of INOPAL, who was involved in the process, considered the success of the necklace with AGROSALUD to be an inspiration for further research. He reported visiting a Georgetown study site in the Guatemalan highlands while developing the new project and finding that IRH had not yet begun to use the necklace in its research. Later, Georgetown and APROVIME collaborated on natural family planning research.

More recently, a Miami-based television crew for the program *Ocurrió Así* traveled to Guatemala to film a story on the necklace method. *Ocurrió Así's* Guatemala correspondent said that they chose the topic because the method eliminates many of the barriers to contraceptive use faced by Guatemalan women and has relevance for women of Latin American origin in many countries, including the United States. Telemundo broadcasts the program throughout North and South America, and in some European countries. Although the story just aired in June 2001, several organizations from El Salvador, Nicaragua and Honduras have contacted the station and expressed a desire to collaborate with APROVIME to offer the method in their counties.

Because the algorithm for integrated services has not enjoyed the same international popularity as the necklace, the links between the Guatemala studies and algorithms adopted in Mexico, Honduras and Bolivia are easier to trace. Regional staff based in Mexico worked with Population Council staff and Ministries of Health in these countries to develop country-specific tools based on the concept of the Guatemalan algorithm. The impact was greatest in Honduras, where the elaboration of the algorithm coincided with the revision of national norms by the MOH. However, each algorithm was developed using a more streamlined process than the three algorithm studies in this review (#17-19), and tailored to the needs of a specific country — for example, the format for the Honduran version entitled *Guía de Atención Integral de la Mujer* (Guide to Integrated Services for Women) is not a list of questions but a flipchart with descriptions of services to provide on the back. Because both the development and testing process and the product varied somewhat from country to country, we gave the algorithm studies a score of "2" rather than "3" on this indicator.

Whether the above were replications of Population Council interventions or new research or activities inspired by previous studies, the vast majority of projects were not used to improve reproductive health services in other countries. The two most likely reasons are the predominantly local dissemination of study results and the unique characteristics of the Mayan populations in the western highlands, for which most of the interventions were designed. It should also be noted that international replication was never an explicit objective to be achieved by the OR projects.

Policy changes

- **In 6 of the 19 studies, there was a change in policy that can be linked to the OR project (I-9).**

The most substantial nationwide policy change attributable to any of these studies has been the inclusion of the algorithm (#19) in the MOH norms.[9] Other changes have been more local (e.g., in the Health Areas of San Marcos and Sololá) or within NGOs or IGSS. Policy changes at these levels have been quite common and generally consist of incorporating the intervention tested into standard operating practice. The first algorithm study also included the introduction of Depo-Provera, with providers trained by Population Council staff. The MOH was so pleased with the outcome that it granted permission to provide Depo only to providers trained by the Population Council. This policy was later changed in response to the growing demand.

Additional funding

- **In 8 of the 22 studies, the original donor funded new or expanded program activities based on the results of the OR study (I-10).**

Several studies led to additional funding, but rarely from new sources. The results of this and previous work suggest that OR results are used to guide the program decisions of groups already receiving funds from a donor (e.g., USAID), but they do not seem to attract new funds from that donor.

Any funding for new activities tended to come from the original source (i.e., the organization implementing the study) or from USAID through Calidad en Salud. Large plantation (*finca)* owners who support AGROSALUD committed larger sums of money following the OR project (#12). The Population Council went on to fund further research or technical assistance, but funding has generally been provided to conduct new OR or to implement or scale up interventions in the short term. The most notable "new funding" that was not provided by USAID or other donors went to health centers in San Marcos. Part of the tripartite strategy (#20) was to work with the local government to get support for the health center. Health centers managed to get funding from the municipality to improve their appearance, for example for painting or signs. More importantly, the improvements to services made through re-engineering enabled the health centers to justify asking for more money from the MOH. While they had previously only been able to execute 13% of the health budget, they could now execute 80%. Funding for equipment and medicine also increased.

- **In 7 of the 22 studies, other donors provided new or expanded funding for program activities based on the results of the OR study (I-11).**

[9] The norms are currently being modified, and numerous key informants said they would now include systematic offering of integrated services through use of the algorithm.

Funding from other donors was unusual, but it did occur. Georgetown has given additional funding to APROVIME (the organization that tested the necklace method in project #11) to do further work in natural methods. AVSC supported continued postabortion family planning in IGSS hospitals (#15). An informant from the Population Council said that AGROSALUD received some funding from a new donor (#12), but was not sure who the donor was. (Current AGROSALUD staff is not familiar enough with the OR project to clarify.)

Despite this handful of examples, the majority of OR studies did not attract funding from new donors. Thus, we have little evidence that OR studies prompt the original donor to invest further in the organization (beyond what was already "budgeted" long-term for the organization to receive) or that they draw new donors to that organization. Moreover, when new funds or a new donor are forthcoming, it is often difficult to determine whether the new resources are in fact attributable to the OR.

IV. DISCUSSION

This review of the OR projects in Guatemala offers an opportunity to reflect not only on the performance of the Population Council in a specific country, but also on the broader questions of what makes for successful outcomes in the conduct of research and in the utilization of results. The experience of this one country prompts us to ask: What *should* be the role of the Population Council (or similar agency providing technical assistance in OR) in assuring the successful completion of this type of project and subsequent changes in program management and policy? We also take advantage of this review to make further observations on the methodology that has been under development for the past two years to evaluate the process and impact of OR studies.

A. Strengths and limitations of Population Council performance in Guatemala

Strengths. The scores from Table 2 represent a crude measure of the strengths and weaknesses of Population Council performance in terms of the conduct of this set of studies (see final column that shows the proportion of projects scoring "3" on each item). Over three-quarters of the projects got the top score of 3 on the following items:

- Research was relevant to local program managers (1.00)
- Results were disseminated to key audiences (.91)
- TA was provided in a sound and collegial manner (.89)
- Implementing organization participated actively in the design of the study (.86)
- Implementing organization participated actively in the conduct of the study (.86)
- Study was completed without delays that would compromise validity (.86)
- Results were judged to be credible and valid (.82)
- Continuity in key personnel was maintained (.77)

On the remaining six process indicators, the proportion of projects obtaining the top score ranged from 0.59 to 0.68. In other words, on every measure of progress at least half the projects in this portfolio received the top score. These scores reflect very favorably on the conduct of this research.

The proportion of studies receiving the top score was considerably lower on the impact indicators, ranging from 0.00 (projects replicated in another country) to 0.82 (the activity tested under the intervention [and shown to be effective] was still observable 36 months later). However, as we will discuss below, the Population Council has far less influence and control over the actual utilization of results than the conduct of the research.

Limitations. One area for improvement in this set of studies concerns the objectives. Some projects had as many as 10 objectives. This effort to "be comprehensive" seemed to set projects up for falling short of achieving all of the stated objectives.

A second problem was "what constituted an objective." In an OR study that tests an intervention, the objectives should clearly state the criteria to be used in judging the

effectiveness of the intervention (e.g., changes in knowledge, attitudes, or behavior among clients; KAP changes among the target population; increases in the number of visits or CYP at a service delivery facility). Thus, an appropriate objective would be "to test the relative effectiveness of strategy A versus strategy B in increasing contraceptive use in district X." Studies often have multiple objectives (but not 10) to capture different types of outputs and outcomes: for example, an improvement in quality of care at the local clinics and an increase in contraceptive prevalence among the target population.

The studies in this portfolio often intermixed the results to be achieved among the target population with the activities to be completed (e.g., training to be carried out, IEC materials to be produced). The latter are important in planning one's strategy to achieve the project's objectives, but they are misplaced as a statement of objectives. (One exception relates to TA projects, in which it is appropriate to consider the completion of activities as a result in itself.)

Why do appropriate objectives matter? Project staff must maintain a clear focus on what they are trying to achieve. Specifically, they should be very clear on how the intervention will be evaluated. Listing the strategies to be used in achieving these objectives (e.g., training, production of IEC materials) provides further clarity to project staff as to what is expected. But the completion of those activities is not an end in itself.

B. Value of conducting OR in a politically hostile environment for family planning

When the Population Council began work in Guatemala in 1988, the social and political climate was very hostile toward family planning. Family planning services for indigenous women were rare, and where they did exist, numerous barriers made them virtually inaccessible. Before the Population Council could begin to concentrate on more specific aspects of reproductive health, they had to legitimize family planning as a health service, and considered OR to be the most effective way to accomplish this goal.

In such hostile environments, OR has several advantages over other approaches. Firstly, organizations are often willing to implement changes within the context of a research project; these otherwise controversial changes may be perceived as less threatening if they are "experimental." Secondly, OR studies are conducted on a small scale, reducing the likelihood of opposition prior to and during the study. Positive results make scaling up interventions more acceptable. The Population Council chose research topics and counterparts that had the potential to improve the general reproductive health environment. According to the Deputy Director of INOPAL, "Our success . . . was that through OR we engaged many stakeholders in a constructive dialogue over a long period of time and thus helped legitimize family planning."

With the USAID/Guatemala Buy-In in 1993 (followed soon afterward by the Cooperative Agreement), the Population Council opened a Guatemala office. They began to do more research, with a broader agenda that included several diagnostic studies and intervention studies involving men, traditional methods and CBDs, among others. The Population

Council worked primarily with NGOs and MOH districts because proposed studies were more easily accepted at these levels. OR projects continued to introduce or improve family planning services within the context of maternal/child health or primary care. Seeing the success of the Population Council studies, the MOH and other organizations became increasingly interested in collaborating with the Population Council. The series of algorithm studies, though carried out in limited areas, was conducted with the central level MOH and, according to a Population Council staff member, showed that family planning is "not a weapon of imperialism but health care."

In a dramatic example of the change in atmosphere, and in the Population Council's relationship with the government, the Population Council and Universidad de San Carlos have developed a distance learning program for senior doctors and nurses from the health districts. The course includes the algorithm for providing integrated services and other Population Council job aids not developed through OR. It is currently being tested in health areas, and will be expanded nationwide once the testing is satisfactorily completed.

Recently, service delivery organizations have become more enthusiastic about OR. Even though few have the capacity to do it on their own, implementing organization staff (including key decisionmakers) are choosing to be involved throughout the OR projects. The next logical step for the Population Council is to build the capacity of these organizations to conduct OR with minimal TA, discussed further below.

C. Appropriate level of technical assistance in OR projects

This review of the Guatemala experience raises the question of the appropriate level of technical assistance for agencies such as the Population Council in the conduct of OR projects. Guatemala is a country that might be classified as "mid-level" in terms of technical capacity for the conduct of applied research. It is far better off than many other developing countries (e.g., many of the African countries), yet it is less well developed than others. In such a situation, what amount of technical assistance is the right amount?

This review revealed that the Population Council played a significant role in the following areas:

- Identifying appropriate topics for OR (although the final selection generally went to the implementing organization);
- Preparing the proposal to meet "international standards";
- Supervising the training of field personnel and quality control of data (although some implementing organizations were better equipped than others to do these tasks themselves);
- Conducting data analysis and interpretation;
- Identifying programmatic implications of the data and preparing recommendations (which counterparts would then review and modify); and
- Preparing the final report.

Should the TA agency be this involved in every stage of the process? On one hand, we applaud the strong technical assistance that the Population Council provided at all stages in the design and conduct of research, thus ensuring a product acceptable by international standards. On the other, we raise the issue: By "doing it all," did the Population Council do all in its power to build technical capacity within these implementing organizations?[10]

On balance, our assessment is that the degree of technical assistance provided by the Population Council (or other technical assistance organizations for OR projects) must be tailored to the local environment. In most cases (APROFAM being the exception), the implementing organizations did not have the technical know-how to carry out research and they could not have participated in an OR study without strong TA from the Population Council. (One key informant had great praise for the patience exhibited by one of the early Population Council advisors in walking the organization through every step of the process over the course of numerous TA visits.) In such circumstances, the Population Council must take the lead if there is to even be a project. At the same time, we point out that the development of local technical capacity in different aspects of the OR skill set should be a by-product of these studies, and considerable work remains to be done on this score. The task is further complicated by the substantial rate of turnover within organizations, though employees often leave one organization only to resurface in another implementing agency.

D. Factors that influence the successful conduct of OR projects

The process indicators in Table 2 reflect the authors' biases as to what contributes to the successful conduct of OR projects. Indeed, this set of projects collectively scored well on these items. During our interviews with key informants, other factors came to light.

OR projects tend to be successful:

- **When a charismatic leader is involved.**

  A charismatic (technical) leader from either the research team or the implementing organization can contribute markedly to the success of a project. In our interviews for this review, key informants of all types consistently attributed the outcomes of a particular study to one individual who inspired those implementing the intervention. When asked about their working relationship with the Population Council, implementing agency staff generally spoke of one or two individuals with great respect and indicated that the study could not have been done without that person. Comments from the Population Council about their counterparts were similar. Describing the commitment of one NGO director, a Population Council informant said that in conducting OR "the people are the most important thing."

[10] Note: this statement overstates the reality and is included simply to raise the issue. Indeed, one can cite several projects in which ARPOFAM hired an external consultant as temporary staff and in fact this individual "did it all" (and went on to become a staff member of PC/Guatemala).

- **When the intervention is closely supervised and monitored.**

Staff from implementing agencies – not Population Council representatives – cited the importance of adequate monitoring and supervision of the intervention itself. In the third algorithm study, "Training in the Algorithm for Integrated MCH Services and Training in Counseling and Family Planning" (#19), the Population Council trained nurses and doctors who in turn trained other providers in use of the algorithm. At the outset some of the trained providers found that offering integrated services was difficult or time-consuming, and without supervision they might well have reverted to their customary way of providing services. However, with Population Council supervision and assistance and the passage of time, providers became more comfortable with the new process and were able to apply the algorithm more easily and efficiently.

In a second case, the community IMCI study (#22) required confirmation that promoters and nurse auxiliaries were diagnosing and treating conditions correctly, but low levels of education made provider records an unreliable data source. For example, promoters were able to count a child's respiratory frequency accurately but sometimes made mistakes in writing it down. Direct observation and supportive supervision were therefore essential in ensuring that diagnoses and treatments were correct.

- **When the tools needed for the intervention are simple.**

The tools used in the intervention must be appropriate for the users: simple enough for local staff to learn and use regularly. The "algorithm" developed to provide integrated RH services in the MOH (#17) actually consisted of two parts: a one-page algorithm with seven questions to ask the client and services to offer according to her responses, and an accompanying manual detailing specific actions to take in carrying out each step. The algorithm itself was quite easy to follow, but the manual was more complicated. With many providers already hesitant to use any job aid in front of clients, many chose not to use one they had difficulty following. The Population Council conducted further OR studies (#18 and #19) to simplify the tools and the training procedure. Only one question was removed from the algorithm, but the manual was considerably revised, eliminating skipping between sections among other things. Providers found the new algorithm and manual much easier to use, and many are now comfortable enough with the questions and the steps to take to provide appropriate services that they can refer to the manual less frequently and still follow the protocols.

The Population Council and their partners developed the community IMCI tools (#22) for a population of providers and clients with little or no formal education, and found they were well accepted. Materials consisting primarily of graphics

helped overcome several potential problems. Rural promoters were able to understand the materials themselves, and were able to use them to educate parents on how to prevent their child from becoming ill in the future. There are many Mayan languages, but all linguistic groups could use the same materials (with the limited text in Spanish). And finally, the pictures reinforced the messages for both the promoters and their clients.

- **When the study design is "do-able."**

  The test of the acceptability and feasibility of using CBD promoters as a vehicle for the distribution of Depo-Provera (#4) was a good example of a very focused project that had clear objectives and a feasible methodology. The project worked well despite the low level of education in Mayan areas.

  In contrast, several projects suffered from being too ambitious. The CBD re-engineering project (# 5) addressed a priority issue for APROFAM: how to restructure its CBD program to be more effective. However, it had three stages, with five interventions at the final stage. Despite an enormous amount of hard work, the team was not able to complete all five interventions. The complexity of the plan of analysis (that broke down the results by salaried versus volunteer workers, men versus women, new staff versus old staff) made it difficult to arrive at concrete conclusions. One Population Council representative commented that it would have been better to break the study into several more manageable components for two reasons: (1) feasibility of conducting the entire study, and (2) attribution of change to a specific intervention.

E. Factors that influence the utilization of OR results

The very purpose of OR is to provide results that will be used to guide improvements in service delivery. In this section we describe factors that influenced the utilization of results in the Guatemala context.

The results of OR studies get utilized:

- **When the intervention is a good match for the implementing agency.**

  Certain interventions tested were an excellent match for the institutional culture of the organization; others were not. This factor had a clear influence on the utilization of results in a number of the projects.

  One of the best "fits" was the testing of the efficacy and acceptability of the necklace (a device to facilitate the use of the rhythm method) among NGOs working in Mayan areas (#11). These NGOs were interested in birth spacing and wanted a method that would be culturally acceptable in Mayan communities. The

necklace method had considerable promise since it is "natural" and does not cause side effects associated with hormonal methods. The OR project demonstrated a high level of continuation at 12 months, indicating that Mayan couples could use the method effectively. As a result, four of the five NGOs that tested the method continued to promote its use after the project, and several other organizations (including the MOH) have since picked up on it. The characteristics of the method fit well with the NGOs' desire to promote a method acceptable to its target population.

The community IMCI project (#22) also fit the needs and experience of NGO health promoters and rural nurse auxiliaries. The providers see many gravely ill children; to the extent they can treat and cure them, they gain the respect of the community. This respect and the increased self-confidence of the providers in turn motivates them to give even better services and establishes the basis of trust for offering potentially controversial services such as reproductive health.

The clinic-based FP/RH algorithm (#17-19) was also a good match for the institutional culture of the organization testing it, the Ministry of Health. The MOH is a large health system and strives to offer standardized care according to established protocols at all service delivery points. By incorporating an algorithm with an accompanying handbook explaining the necessary steps in greater detail, the MOH seeks to ensure that providers in all centers offer the same health services in the same way to women with a specific need. While flexibility is desirable in other interventions, here the implementing organization sought consistency.

The case of AGES drives home this point of the importance of "good fit" between the intervention and the organization. AGES conducted two successful studies (#7 and #8) involving health education strategies to reach indigenous populations. They felt comfortable in applying these strategies and carried them out with minimal problems. However, in a third study AGES, under some pressure from the INOPAL Director, included a contraceptive service delivery component. Distributing contraception was not one of AGES' competitive advantages; in fact, AGES did not want any of the "bad press" that APROFAM was getting for doing this type of work. When problems arose at the field level AGES staff, inexperienced in dealing with such situations, retreated from this aspect of the intervention. Service delivery, while an understandable goal, is not part of their mission, and the project floundered when they attempted to integrate contraceptive services into the project design.

An IGSS study (#16) also underscores this point. The OR project attempted to improve rural perinatal care by training TBAs, but IGSS' institutional strength is in clinical service delivery with clinically trained personnel. They were therefore not willing to invest in training all affiliated TBAs when they could spend the money in hospitals that delivered far more births than all the TBAs combined, even though the TBAs reached different groups.

- **When the intervention yields immediate, observable improvements.**

  Organizations are also more likely to utilize research findings when an intervention yields immediate, observable improvements. The re-engineering activity in MOH centers in San Marcos (#21) took only a few weeks but resulted in drastic changes to clinic appearance and atmosphere. Both providers and clients noticed the new paint on the walls, new signs and cleaner facilities, as well as the shorter waiting time. Nurses were better utilized so they experienced greater job satisfaction and the clinic was able to attend all patients arriving in a given day when previously many had been turned away. Auxiliary nurses participating in the continuing education course (#20) had greater capacity to attend and counsel clients, or understand the doctor's or nurse's diagnosis and treatment in a clinic setting. In contrast, resistance to interventions has been strongest when benefits are less obvious. Doctors who were trained but did not use the algorithm (#17-19) said that it took too much time and that they didn't notice any negative consequences of not offering women services they might accept but have never had.

- **When the intervention increases provider motivation.**

  Providers feel motivated when they are able to offer clients a better quality of care. Several providers mentioned that they were willing to do more work than they had in the past because they felt more competent and their patients were more satisfied. They referred to their experience with the Population Council as "teamwork" and said they wanted to continue the intervention and were more open to changes or new interventions than they had been before the OR study.

- **When the intervention has a committed advocate.**

  Much as the "charismatic leader" contributes to a quality OR study, a "champion" or advocate for the study who remains committed after it ends enhances the likelihood that the intervention will be scaled up or replicated and institutionalized**.** One of the top-level officials in the Reproductive Health Unit of the Ministry of Public Health was an advocate for continuing to test and improve the algorithm, and it eventually became a standard protocol for MOH hospitals, health centers and health posts (#17 – 19). A locally hired Population Council representative remained very involved in the scaling up and replication of the re-engineering component in San Marcos and Sololá (#21).

  An extension of the individual advocate is a committed organization. At both the administrative and the service delivery levels, the MOH believed that initial problems with the algorithm could be solved, and they invested several years and

considerable human and financial resources in developing and implementing a satisfactory algorithm. Had their commitment dwindled at the end of the first study, the algorithm might never have been developed into a useful tool and indeed could have been abandoned.

The commitment of the MOH to using OR to improve services did not end with the algorithm studies. After seeing the benefits of an OR intervention, providers and decision-makers were more willing to participate in subsequent studies. Positive experiences with one OR study often led to more openness to future interventions in NGOs as well. In addition to participating in the IMCI study (#22), both CDRO and PIES del Occidente trained their clinical staff to provide integrated care through the RH algorithm. PIES del Occidente now requires that new doctors hired be trained and experienced in the algorithm prior to working at PIES clinics.

- When technical assistance continues beyond the end of the project.

  Implementing organizations may be more likely to integrate OR findings and scale them up within the organization if they receive some technical assistance and monitoring throughout the transition process. A representative from an NGO that only recently began offering reproductive health services said, "Things improved when the [Population Council] Xela office opened." A nurse from the health center in San Marcos said, "We think we are doing things right, but we need feedback to be certain."

- **When the timing is right.**

  A major factor in all OR studies is the timing of the intervention. Indeed, it may be difficult to distinguish the impact of an OR study from an idea whose time has come. Although the quality of the study is important, other factors may be responsible for scaling up or replication of the intervention in question. For example, this portfolio of projects contains several studies in which Depo-Provera was introduced in a particular service environment and met with unprecedented demand. APROFAM CBDs were trained to provide Depo as part of study #4; MOH doctors and nurses, and later IGSS doctors and nurses, received training in Depo as part of studies #18 and #16.

  Can these different OR studies claim credit for the fact that Depo-Provera is currently the most widely sought method among new users? Without a doubt, the OR projects contributed to further increasing awareness of the method and improving access to quality services for its provision. However, one could also argue that Depo, already a popular method elsewhere in Latin America, was poised to become an important method in Guatemala. It is simple to use and highly effective; it can be used without the knowledge and cooperation of the

partner; and it appeals to the Guatemalan affinity for injections as the most effective form of medication.

F. The value of evaluating diagnostic studies with this methodology

In previous case studies using this evaluation methodology, evaluators have chosen to include only intervention and evaluative studies, and to exclude diagnostic studies.[11] Although the original scope of work for this review excluded diagnostic studies, the Population Council requested that they be included. Implicit in this request was the belief that diagnostic studies **should** "lead to something"; otherwise, what is the value of doing them? Thus, in this review we examined whether the two diagnostic studies[12] led to new interventions or changes in service delivery. These studies were:

- Baseline Study of Reproductive Health Beliefs and Attitudes of Males in Four Health Districts in the Department of El Quiché (#2); and

- Study of Cognition and Speech Patterns of Urban and Rural Indigenous Community Residents about Reproductive Health in the Department of Quetzaltenango (#13).

Of these two diagnostic studies, one led to an intervention. The baseline survey for reproductive health activities among men in El Quiché (#2) served to inform the intervention that was tested immediately afterwards (#3). Although the actual intervention was one of the least successful in this portfolio of studies, the timing and use of the diagnostic study were highly appropriate. In this case, one might argue that the two studies did not need to be separate; rather, the diagnostic component should have been (and was) the logical first phase of the larger project. However, there was value in making this activity into two separate projects. Had the diagnostic study provided results that suggested **against** implementing an intervention, one would have been spared the effort and expense of the implementation phase.

The second diagnostic study may have informed the work of the Population Council in Guatemala, but it was not clearly linked with an intervention. In 1994 the Population Council commissioned the Universidad del Valle to conduct an anthropologic study on cognition and speech patterns regarding reproductive health (#13) to inform the design of subsequent reproductive health projects for rural Mayans. The research showed, among other things, the impracticalities of condom use and low levels of knowledge about reproductive physiology and pregnancy, which helped explain attitudes toward contraception. However, there was no clear dissemination strategy for these study results and due to a change in Population Council leadership in August 1996 and a shift in

[11] Many of the indicators do not lend themselves to diagnostic studies, and it was decided that this methodology was not appropriate for these studies.

[12] Several other studies included diagnostic components, but we classified them as "demonstration/evaluative" or "intervention" if they also involved some aspect of actual service delivery or a related intervention (e.g., sex education in Mayan communities).

priorities, the results did not reach the intended users. Although the findings were summarized in two Population Council publications,[13] and have the potential to benefit the recent work of NGOs, the results remain underutilized by any measure.

Our conclusions from examining these diagnostic studies are as follows:

> (1) This methodology can be used in evaluating diagnostic studies, although a number of the indicators for impact are by definition "not applicable."
>
> (2) Studies that deal with "new areas" are often mislabeled as "diagnostic" when in fact they are demonstration/evaluative. To avoid missing such cases, evaluators should assess "diagnostic" studies as well.
>
> (3) Diagnostic studies should contribute to interventions, and thus they should be included in future evaluations of this type.

G. Reflections on the methodology

This review brought up a number of issues related to the application of the methodology. Several issues are presented below while those related to dissemination are included in Appendix D, which presents the results of pretesting three new questions on dissemination.

➢ Assessing the "whole" through the lens of the individual studies

This evaluation was based on a systematic review of the 22 projects that comprised the portfolio of OR conducted in Guatemala with assistance from the Population Council from 1988-2000. We interviewed key informants about the details of each specific study, and we compared their responses to the findings published in the written reports. However, the methodology used a "case-by-case" approach. We did give key informants the opportunity to explain contextual factors that facilitated or hindered the conduct of the research and the utilization of results, but most discussion still focused on individual projects. Although we attempted to look at the projects in the context of the overall political and social context of reproductive health interventions in Guatemala during this period, the line of questioning was not designed to capture the "bigger picture" and we may not have done it justice in this review.

In short, one can criticize this methodology for focusing on the "trees," and losing sight of the "forest." In future applications of the methodology, the evaluators should address how to overcome or minimize this limitation.

---

[13] Population Council/Guatemala. 1998. *Findings and Lessons Learned In Delivery of Reproductive Health Care to the Rural Mayan Population of Guatemala from Operations Research and Diagnostic Studies 1994-1997*. New York: Population Council; and Enge, Kjell. 1998. *Salud y reproducción: qué piensan, sienten y desean los mayas*. Documentos de Trabajo, núm. 20, 1998. Mexico: INOPAL III.

- **Recall**

As in the previous six case studies, we found that key informants had more difficulty answering questions on projects completed long ago than those completed more recently. Some of the studies in this review were completed over ten years ago, and even researchers had some trouble remembering what had happened in the study itself, or had lost contact with the implementing organization and could not adequately respond with respect to the impact indicators. In these situations, key informants either had to guess, refer to the final report, or not answer. Using the final report to answer questions on the project reduces the value of interviewing a key informant, because one of the main purposes of the interview is to learn more about topics that are *not* in the report.

Recall was particularly hazy in relation to dissemination activities, levels of participation of both the Population Council and the implementing organization in the design and conduct of the study, and occasionally even the study findings, all of which are important factors for enhancing the likelihood of utilization.

In this review we were able to get some idea of the process and impact of all studies, often by consulting additional sources, but the data would have been richer if all projects were fresh in the minds of the individuals interviewed. Our experience with this Guatemala review supports the recently established approach to evaluating FRONTIERS projects, which consists of a process assessment at the project end and an impact assessment three years later (when presumably key informants will have some recall of the experience of each study).

- **Variety of informants**

In some cases a single key informant and the final report provided sufficient data to answer all of the indicators on the data collection form (see Appendix C). However, the team found it necessary to interview several key informants per study, because people tended to remember different things or perceive things differently. The FRONTIERS Evaluation Guidelines[14] recommend interviewing key informants from three main categories: 1) researchers, 2) program managers, and 3) donor agency staff, "ideally one from each category." Experience in this review suggests that the *minimum* should be one from each category, including one Population Council researcher as well as his or her implementing agency counterpart. Without this range of experience, it is difficult for key informants to provide information on the full range of indicators (e.g., collaboration and participation, credibility and validity of study findings, feasibility, availability of research results and impact on the original implementing agency and other organizations). Indeed, we found that interviews with two different individuals within the same organization (either at the service delivery or research level) could yield either complementary or conflicting information.

[14] FRONTIERS/Tulane University. *Evaluating Operations Research Utilization: Guidelines for Assessing Process and Impact.* Washington DC: Frontiers in Reproductive Health, November 2000.

- **Impact assessment by Project Monitors**

Currently, evaluation under FRONTIERS consists of a Project Monitor completing the Process and Impact Assessment Forms at the appropriate times, followed by a verification visit by FRONTIERS/Tulane staff to a subset of projects. Interviews with key informants may be used as needed, but are not required based on the Project Monitor's presumed knowledge of what has happened in terms of changes in service delivery and policy after the study was completed (i.e., "impact"). This review showed that the Population Council Monitor or PI often knows what has happened only to the extent of his or her involvement. Many may continue to be aware of new activities related to the study, but usually the implementing agency and the end users (if these are different) know better what has happened and why. If they have made modifications to the intervention adopted, they may not inform the Population Council. Project Monitors will likely know what parts of the intervention were officially adopted, expanded or replicated, but will probably not routinely follow up on these actions with the end users to determine the extent of institutionalization or program level reactions to the changes. In the intervening three years from the project end to the impact assessment, a Project Monitor will be involved in a variety of new studies, possibly with different organizations; thus, he/she may not be able monitor completed projects very closely. In other words, Project Monitors could conduct an impact assessment based on their own knowledge, but to get in-depth information on the context and nature of the impacts, they should include key informant interviews as part of the evaluation process.

- Issues with specific questions

In addition to the above reflections on the methodology as a whole, we found that some specific indicators posed a challenge in this review. Four indicators are listed below, along with a description of the problem and a proposed modification that might resolve the problem.

Indicator P-8: *The implementing agency judged that the technical assistance was useful, methodologically sound, and provided in a collegial manner*. Responses to this indicator were very nearly all extremely positive, even though the interviewers were not Population Council employees, and PC/Guatemala staff were not present. This apparent consistency suggests the possibility of a courtesy bias on the part of the implementing agency staff, many of whom continue to work with the Population Council on other projects. The enthusiasm of the responses, on the other hand, as well as references to specific experiences, leads us to believe that most of the responses are genuine and do not reflect a generalized bias. Nevertheless, an outside evaluator (not a Population Council staff member) must ask this question, and responses must be interpreted with caution. Since a Project Monitor should not be asking the implementing agency this question, we will need to either only ask it in verification visits by Tulane staff, or develop a supplementary form to be returned directly to the Washington, DC office to provide

respondents with greater confidentiality. Without these measures, few organizations can be expected to be candid in their assessment of technical assistance or even active participation or collaboration with the Population Council.

Indicator P-9: *The study design was methodologically sound.* We decided to use a fairly strict interpretation of "methodologically sound" in this review. A number of the studies produced seemingly valid results but did not use experimental or quasi-experimental designs typically associated with OR. Designs such as pre and post-test with no comparison group, or post-test only with or without a control group kept the studies simple and feasible, yet these designs might not meet international criteria for "rigor." In instances where we felt studies could have been more "scientific" in their designs but nevertheless produced credible and defensible results, we have not given the studies the top score of "3." In future applications of the methodology, it will be important to ensure that scoring criteria are consistent.

Indicators I-10 and I-11: *The original donor funded new or expanded program activities based on the results of the OR study;* and *Other donors provided new or expanded funding based on the results of the OR study.* We found the term "new donor" difficult to define. Many reproductive health programs in Guatemala, as in the rest of the world, are funded by USAID but have different objectives and activities. Calidad is the most evident of this type of "new funding" in Guatemala. Are funds provided by another USAID cooperating agency from the same or "different" donors? On the one hand, a single donor, USAID, funds Calidad and the implementing agencies, but on the other hand, each program operates independently and decides to allocate funds based on its own priorities and through its own decisionmaking process. In fact, only very rarely did any non-USAID CA provide new funds. One option is to reword these indicators to make them more precise, while another is to include all direct or indirect USAID contributions as the original donor, with the evaluator specifying the exact source of the funds.

Table 1.
OR Studies in Guatemala by Type and Substantive Focus

	Type of Study					
Topic of Study	Technical Assistance	Diagnostic	Demonstration/ Evaluative	Intervention	Dissemination/ Utilization	**Total**
Acceptability		2				**2**
Access			2	1		**3**
Capacity Building				2		**2**
CBD workers				3		**3**
Community promotion			1			**1**
Constellation of services			1	1		**2**
Institutionalization					1	**1**
Integration			1	2	3	**6**
Job instruments				1		**1**
KAP			2			**2**
Men		1		1		**2**
Monitoring and supervision				1		**1**
PAC			1			**1**
Service organization			2			**2**
Sustainability		2				**2**
Technical competence	2		1	1		**4**
TOTAL	**2**	**5**	**11**	**13**	**4**	**35**

Table 2.
Summary of Key Findings from the 22 OR Studies and Subsequent Changes in Service Delivery or Policy

Study	**Type of OR Study**	**Main Finding of the Study**	**Organization Acted on Results (I-2)**	**Evidence 36 Months Later (I-3)**
#1. On-site Training: CBD Promoters in Indigenous Areas	Intervention	Phased, competency-based supervision improves distributor knowledge.	APROFAM switched from simply resupplying promoters and collecting money to a system of "*supervisión capacitante*" and new scheduling of supervision.	This system remained in place thereafter.
#2. Reproductive Health Beliefs and Attitudes of Males in El Quiché	Diagnostic	Men had very low knowledge of family planning but they were interested in learning about birth spacing.	APROFAM planned and executed a project on "Appropriate Health Education Strategies for Men in 4 Health Districts in el Quiché." (see next project)	N/A
#3. Appropriate Health Education Strategies for Men in El Quiché	Intervention	Recruiting men in El Quiché to attend health education talks didn't work; instead, they were willing to attend recreational and athletic events. Contraceptive prevalence was higher in experimental than comparison communities.	No immediate follow-up	APROFAM has recently initiated new activities directed to men, but with no apparent link to the original study.

Study	**Type of OR Study**	**Main Finding of the Study**	**Organization Acted on Results (I-2)**	**Evidence 36 Months Later (I-3)**
#4. Injectable Contraceptives Provided by Volunteer Community Promoters	Intervention	Trained community personnel can effectively and acceptably provide Depo-Provera to Mayan communities.		APROFAM continues to authorize selected CBD promoters to provide Depo at the community level.
#5. Re-engineering CBD Program of APROFAM	Demonstration/ Evaluative	Identified profile of "ideal CBD worker" from perspective of Mayan women. Training promoters in broader reproductive health topics increased sales.	APROFAM developed new criteria for selection of promoters.	
#6. Improving Client Referrals to APROFAM Clinics	Intervention	The strategy to bring promoters to be trained in the clinic was more effective in increasing referrals to APROFAM clinics than the usual strategy of training by *educadores*, but was more expensive.	Intervention not continued due to cost.	N/A
#7. Interest among Indigenous Populations in Learning about Family Planning	Demonstration/ Evaluative	Women in Mayan communities were interested in learning more about reproduction and birth spacing. It was acceptable to provide health education talks in Mayan communities.	AGES continued to develop educational programs for Mayans.	This work continued through the 1990s.

Study	**Type of OR Study**	**Main Finding of the Study**	**Organization Acted on Results (I-2)**	**Evidence 36 Months Later (I-3)**
#8. Reproductive Health Education in Indigenous Areas through Bilingual Teachers	Demonstration/ Evaluative	Bilingual teachers can be motivated to effectively teach reproductive health subjects in Mayan communities at a cost of $2.50/student.	AGES continued its work with sex education in Mayan communities, with additional support from other donors.	AGES remained active in this area three years later (though recently had funding cuts from the government).
#9. Access to Reproductive Health Services and Education in Indigenous Communities	Intervention	AGES was strong in conducting educational activities, but the service component was not successful: few teachers distributed contraceptives, segmentation meetings didn't work, and two other strategies were abandoned.	AGES remained very reticent to work in the delivery of contraceptive services.	N/A
#10. Providers and Reproductive Health Service Delivery Strategies in 2 Conservative Indigenous Communities, Lake Atitlán	Demonstration/ Evaluative	Technical assistance helped increase knowledge and skills of workers in Rxiin Tnamet for FP/RH service delivery in Mayan communities; community workers performed better than clinical staff.	Rxiin Tnamet put into practice training and suggestions, improved its MIS system for low literate promoters, and strengthened service delivery in the community.	Rxiin Tnamet remains a model for service delivery in Mayan areas.
#11. Testing the "Blanket Rule" Rhythm Method among Indigenous Guatemalans	Intervention	Mayan couples could effectively use the "necklace method" to prevent pregnancy, and it was highly acceptable in the local culture.	Four of five NGOs continued to promote the method as part of their health services.	These NGOs, as well as the MOH and others, continue to promote the necklace method.

Study	**Type of OR Study**	**Main Finding of the Study**	**Organization Acted on Results (I-2)**	**Evidence 36 Months Later (I-3)**
#12. Self-Financed Incorporation of Family Planning in Rural Fincas	Demonstration/ Evaluative	Reproductive health services are in demand on fincas and can be provided by promoters in health posts and referrals to APROFAM.	Reproductive health services, particularly counseling, continued to be offered at finca health posts.	Services still offered 9 years after study end.
#13. Cognition and Speech Patterns about Reproductive Health of Urban and Rural Indigenous Community Residents	Diagnostic	Reproductive health programs based on Western concepts of reproductive health and knowledge are unlikely to succeed; programs should be planned with knowledge of the culture and lifestyle of indigenous groups.	None (no intervention)	N/A
#14. Collaboration between Two NGOs, ATI and APROFAM, in Delivery of Family Planning Services	Intervention (only diagnostic phase implemented)	Information about barriers to women attending reproductive health meetings and the characteristics of their ideal promoter.	None (no intervention)	N/A
#15. Quality of Services for Women Who Had an Abortion	Demonstration/ Evaluative	Many women not using family planning will use a method if services are more convenient (in particular if they can get a method right away).	Postpartum/postabortion family planning services were continued in OB-GYN hospital and expanded to all hospitals in the IGSS system.	Changes still in effect in first hospital after 3 years; changes in additional hospitals are more recent but still in effect.

Study	**Type of OR Study**	**Main Finding of the Study**	**Organization Acted on Results (I-2)**	**Evidence 36 Months Later (I-3)**
#16. Integrated Obstetric, Family Planning and STD Training for TBAs	Intervention	Self-esteem should be incorporated into training, especially for community workers. Supportive supervision is most effective for reinforcing training/knowledge.	Supervision system was maintained and used in other IGSS services. TBA training was continued on a very limited scale, but was not cost-effective enough to expand.	Supervision system still in use.
#17. Systematic Offering of Family Planning/ Reproductive Health Services in Guatemala	Intervention	The algorithm is a potentially useful job aid for screening clients' reproductive needs, but can only be successful with greater commitment of area and district chiefs.	Further OR studies were conducted to test a modified algorithm and new training strategies.	New model of integrated services for women still in place after 4 years, but has been modified.
#18. Institutionalization of Systematic Offering of Integrated Reproductive Health Services in Quetzaltenango	Demonstration/ Evaluative	Training all medical and non-medical staff in their own work environment and giving providers immediate feedback on use of the algorithm leads to a greater degree of institutionalization than centralized training.	Scaling up has continued using a new training approach. MOH plans to include algorithm in norms for all health centers and posts.	Study completed in 1999. Changes are still in place.

Study	**Type of OR Study**	**Main Finding of the Study**	**Organization Acted on Results (I-2)**	**Evidence 36 Months Later (I-3)**
#19. Evaluating Training in Algorithm for Integrated MCH Services and Training in Counseling and Family Planning	Intervention	Offering services systematically through a standard set of questions asked of all women regardless of reason for visit improves quality of care and client satisfaction (more than family planning training alone).	Algorithm continued to be scaled up, and a new training strategy was developed to promote institutionalization.	Application of algorithm was scaled up, and maintained in original health centers. Study was completed in 1999 and changes are still in place.
#20. Continuing Education System in Reproductive Health for Auxiliary Nurses	Demonstration/ Evaluative	Distance education can improve the knowledge and skills of auxiliary nurses.	Distance education course is offered to new auxiliary nurses entering the system, and the Nursing School is working on a final module with JHPIEGO.	Study completed in 1999. Would like to scale up after Module 6 completed but don't know if they can find funding.
#21. Testing a Tripartite Strategy in San Marcos	Demonstration/ Evaluative	When they feel they have the ability to make changes and can see the benefits, health center personnel are willing to work harder to provide quality care, and satisfaction of both clients and providers increases.	Sololá Health Area adopted intervention and more health centers in San Marcos carried out re-engineering.	In study intervention sites, new model of attention is still in place, with periodic evaluations and improvements. Other health centers adopted the intervention since 1999.
#22. Algorithm for IMCI at Community Level	Intervention	Community workers with low education level can syndromically diagnose and treat life-threatening childhood illnesses using algorithm and IEC materials.	Non-clinical health workers from MOH as well as NGOs began using the IMCI tools to treat sick children.	Study ended in 2000. NGOs and MOH are continuing to expand implementation.

Appendix 1. Hospitals and maternity services available in Guatemala, 2001

Region	Type of hospital					
	Regional	**Area**	**District**	**Municipal Maternity**	**IGSS Hospital**	IGSS Medical office
Nor-westl	Coatepeque Quetzaltenango	Chimaltenango Huehuetenango Sacatepéquez San Marcos Sololá Totonicapán	Malacatán San Pedro Necta	San Carlos Sija		San Lucas Tolimán Sololá Totonicapán
Sur-occidental		Escuintla Jutiapa Retalhuleu Santa Rosa Suchitepequez	Asunción Mita Atescatempa San Rafael Tiquisate		Escuintla Mazatenango Patulul Santa Lucia Cotzimalguapa Tiquisate	La Gomera
Centro-oriental	General San Juan de Dios Roosevelt	Chiquimula El Progreso Jalapa Zacapa	Cabañas Gualán Amatitlán Mataquescuintla	C/S Guatemala Norte C/S Guatemala Sur	Hospital de Gineco-Obstetricia Juan José Arevalo	Chiquimula Jalapa Zacapa
Nor-oriental		Alta Verapaz Baja Verapaz Izabal Poptún Quiché San Benito Petén Sayaxché	Cahabón Carchá El Estor Fray Bartolomé Ixcán Joyabaj La Tinta Livingston Melchor de Mencos Morales Nebaj San Cristobal Senahú Tucurú Uspantán			Salama Sta. Cruz del Quiche